IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of

STRACHAN et al — Atty. Ref.: 922-169

Serial No. 10/721,866 — TC/A.U.: 2661

Filed: November 26, 2003 — Examiner:

For: PACKET-SWITCHED NETWORK AND NETWORK SWITCHES HAVING A NETWORK LAYER FORWARDING ACTION PERFORMED BY DATA LINK SWITCHING

* * * * * * * * * * * *

March 15, 2004

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

SUBMISSION OF PRIORITY DOCUMENTS

It is respectfully requested that this application be given the benefit of the foreign filing date under the provisions of 35 U.S.C. §119 of the following, a certified copy of which is submitted herewith:

Application No.	Country of Origin	Filed
0227614.5	Great Britain	27 November 2002

Respectfully submitted,

NIXON & VANDERHYE P.C.

By: Larry S. Nixon

Larry S. Nixon
Reg. No. 25,640

LSN:vc
1100 North Glebe Road, 8th Floor
Arlington, VA 22201-4714
Telephone: (703) 816-4000
Facsimile: (703) 816-4100

USSN
10/721,866

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation and Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein together with the Statement of inventorship and of right to grant of a Patent (Form 7/77), which was subsequently filed.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed Stephen Hordley

Dated 18 February 2004

Patents Form No. 1/77

The Comptroller
The Patent Office
Cardiff Road
Newport
Gwent NP10 8QQ

REQUEST FOR THE GRANT OF A PATENT

The grant of a patent is requested by the undersigned on the basis of the present application:

1. Title of the invention: **PACKET-SWITCHED NETWORK AND NETWORK SWITCHES HAVING A NETWORK LAYER FORWARDING ACTION PERFORMED BY DATA LINK SWITCHING**

2. Applicants details: 0227614.5

First or only applicant: **3Com Corporation**

ADP: 06396741001

Country: US State: Delaware

Address: 5400 Bayfront Plaza
Santa Clara
California 95052-8145
United States of America

3. Name of agent: Bowles Horton ADP Code: 0008805003

Agent's address: Felden House, Dower Mews
High Street, Berkhamsted
Hertfordshire HP4 2BL

4. Agent's reference: **105625**

5. The application claims an earlier date under any of Sections 8(3), 12(6), 15(4) or 37(4): NO

6. Declaration of Priority (if any)

Country	Application Number	Priority date
None		

Patents Form 1/77 (continued)

7. Inventorship

The applicant(s) is/are the sole inventor or joint inventors: NO

8. Check List of Documents:

A. The application is accompanied by the following number of sheets:

Request:	1
Description:	10 ✓
Claims:	2 ✓
Abstract:	1 ✓
Drawings	5 + 5 ✓ (x2)

B. The application as filed is accompanied by the following:

Patents Form 7/77:	YES ✓
Patents Form 9/77:	YES ✓
Patents Form 10/77:	YES ✓
Priority Document:	NO
Translation of priority document:	NO

9 Signature

---------- for BOWLES HORTON

Patents Form No. 7/77

The Comptroller
The Patent Office
Cardiff Road
Newport
Gwent NO10 8QQ

THE PATENT OFFICE
M
27 NOV 2002
NEWPORT

27 NOV 2002

STATEMENT OF INVENTORSHIP AND OF RIGHT TO THE GRANT OF A PATENT

Patent application No.(new) 0227614.5

Title: **PACKET-SWITCHED NETWORK AND NETWORK SWITCHES HAVING A NETWORK LAYER FORWARDING ACTION PERFORMED BY DATA LINK SWITCHING**

I/We: 3Com Corporation

the applicants in respect of the above-mentioned application for a patent declare as follows:

(i) I/We believe that the person(s) whose name(s) and address(es) are stated on the reverse side of this form is/are the inventor(s) of the invention in respect of which the patent application is made;

(ii) the derivation of my/our right to be granted a patent upon the said application is as follows:

by virtue of the employment of the inventors

(iii) I/We consent to the publication of the details contained herein to each of the inventors named on the reverse side of this form

Signature: ---------- 26 November 2002
for BOWLES HORTON (Authorised Agents)

Telephone: ~~0~~1442 875961

Our reference: 105625

STRACHAN, Graham
10 Southwold Close
Aylesbury
Buckinghamshire
HP21 4EZ/GB 8514903001

MORAN, Paul James
The Bye, 120 Green End Road
Hemel Hempstead
Hertfordshire
HP1 1RT/GB 7174600001

CAPON, David
2 St Margarets Way
Leverstock Green
Hemel Hempstead
Hertfordshire
HP2 4PA/GB 8514911001

STUBLEY, John
105 The Park
Redbourn
St Albans
Hertfordshire
AL3 7LT/GB 8514937001

TD

PACKET-SWITCHED NETWORK AND NETWORK SWITCHES HAVING A NETWORK LAYER FORWARDING ACTION PERFORMED BY DATA LINK SWITCHING

Field of the Invention

This invention relates to packet-switched communication networks, particularly Ethernet-type networks. The invention more particularly relates to achieving efficiency of operation in a complex network, such as one having a multiplicity of subnets or virtual local area networks, and more particularly to the reduction of unnecessary traffic between a core router and an edge switch wherein an edge switch is required to switch packets both between different entities on the same subnet but also between entities on different subnets.

Background to the Invention

Broadly speaking, apart from the media employed for the conveyance of data between devices, data networks are composed of data terminal equipment (DTE) which constitute the sources and ultimate destinations of data on the network, and switching devices which perform, as explained further below, both switching and routing, and which fall into generally two categories, namely edge devices and core devices. Herein, 'edge device' is intended to mean a switching or routing device which is the first encountered by packets on dispatch from data terminal equipment and/or the last encountered by a packet before it reaches its ultimate data terminal equipment. Herein 'core device' is intended to refer to a switching or routing device which is separated from data terminal equipment by an edge device.

Packet switching between members of the same subnet or virtual local area network is commonly performed at the data link or media access control (MAC) level, often called 'layer 2' switching because the relevant layer in the open system's interconnection (OSI) model is the second or data link layer. Switching at this layer is

normally between members of the same subnet, and only the layer 2 (MAC) address information in a packet is required.

Data packets of the kind employed in the present invention will normally have a format that includes a MAC address header, comprising a MAC source address (identifying the device from which the packet has come) and a MAC destination address (indicating the device to which the packet should be forwarded). They will also include a layer 3 or IP header which typically includes an IP or network source address and a network destination address. As these names imply, MAC addresses are used to determine the device to which a packet should be sent whereas a network address identifies the network to which the packet should be sent.

As indicated above, layer 2 switching, normally performed between members of the same subnet, does not normally require any change in the header data of a packet. When a switching device receives a packet, it will perform a look-up in a 'layer 2' database which will contain an entry including the relevant destination address, and (for example) the port forwarding data, typically the number of the port from which the packet should be forwarded to reach that destination of the same subnet. It may also have a field which identifies that subnet. However, routing between different subnets is a more complex activity and usually requires recourse to a routing table which apart from the network destination address will include an identification of the relevant subnet and a MAC address which will have to be applied to the packet. Routers commonly also perform various other functions which are not directly relevant to the present invention.

Summary of the Invention

In a layer 3 network, that is to say a network having a multiplicity of subnets and requiring layer 3 switching, all traffic between subnets will normally travel from the edge of the network and into a core where it will be routed and sent out again to the edge of the network. In some cases the source and destination stations might be connected to the same layer 2 edge device.

The basis of the present invention is the avoidance of an unnecessary return journey of a packet between a layer 2 edge device and the layer 3 core, thereby conserving both up-link and core bandwidth. The edge device can be provided with sufficient addresses to be able to forward the packet by means of a layer 3 look-up if both the source and destination end stations are attached to it but to switch the packet at layer 2 up to the layer 3 core if they are not

Further objects and features of the present invention will be apparent from the following detailed description with reference to the drawings.

Brief Description of the Drawings

Figure 1 is a simplified schematic diagram of a switch.

Figure 2 is a schematic diagram of a router.

Figure 3 is a diagram illustrating a cascade connected switch.

Figure 4 illustrates a fragment of a network.

Figure 5 illustrates a known form of edge switching.

Figure 6 illustrates one switching process according to the invention.

Figure 7 illustrates the fragmentary network of Figure 4 operated according to the present invention.

Description of the Preferred Embodiments

Figure 1 of the drawings is a simplified schematic representation of an edge device which can be used in the present invention. For the most part, the device 10 shown in Figure 1 is of known form; however, as will be seen, it is necessary according to the invention that the device 1 should be capable of switching at layer 2 (employing media access control address data) as well as switching at layer 3 (utilising network address data). Also, the look-ups are interlinked, either in hardware or software as will be described later. Although switches which can switch according to both layer 2 and layer 3 information are known, in essence switches of that nature will attempt to switch at layer 2, and on inability to switch at layer 2 will attempt to switch at layer 3, and their operation will differ thereafter. The difference between the switch in Figure 1 and known switches will become more apparent after a discussion of Figures 5 to 7.

In order to provide a general view of the organisation of the switch, there follows a brief description of Figure 1. The switch 10 in Figure 1 will have a multiplicity of ports, herein represented as merely four ports (instead of a typical number such as twelve or twenty-four). Each of the ports is connected to a port ASIC, which will perform initial and final processing on packets and typically contains the physical sub-layer and data link sub-layer (or MAC). A system of buses is represented in Figure 1 merely by a bus 15. Typically, packets received by any of the ports 11 to 14 will be stored in memory 16 while the headers of the packets are processed in order, for example, to perform look-ups with the aid of look-up databases 17 which can be accessed by a look-up engine 18. For convenience this engine 18 is shown as comprising a layer 2 look-up engine 18a (L2 LU) and a layer 3 look-up engine 18b (L3 LU). The engine 18a will have recourse to a layer 2 look-up table 17a, containing entries accessed by media access control addresses and yielding forwarding information such as port numbers, whereas the engine 18b will have recourse to layer 3 (routing) tables 17b and 17c, containing entries of network addesses and corresponding forwarding information, i.e known routes and possible default routes.

As will become apparent some embodiments will need to preserve a look-up result from the layer 2 look-up even though a layer 3 look-up is performed.

The device includes a processor represented by a CPU 19.

The database or databases 17 (whether the address and forwarding data in the database is in one table or split into a number of tables 17a to 17b is not important) contains certain type of information which will be more particularly described after a discussion of Figure 5.

A switch of the kind shown in Figure 1 is represented in practice by a switch type 4400 made by 3Com Corporation. Such a switch is 'stackable' in that it can be put into a cascade connection with other (similar) switches to form a single switch entity. One purpose of this is to provide a switch with a larger number of ports that a single switch in a simple manner not requiring reorganisation of the network generally.

If the switch is stacked, as for example as shown in Figure 3, packets received at an ordinary port of the switch but intended for dispatch by a port on another unit in the cascade will travel by way of a cascade connection. A variety of ways of organising a cascade connection are known. Figure 3 shows one example. The stacked switch 30 in Figure 3 is a stack of three units, identified as Unit 0, Unit 1 and Unit 2. Each of these units has a multiplicity of 'front panel' ports 31 (corresponding to ports 11 to 14 in Figure 1), a communication core 32 which performs the main functions of the switch, a management agent 27 (which can be controlled by way of management packets coupled to one of the front panel ports or otherwise) and look-up tables 29. Each of the units includes a cascade module 33 which has a cascade port 22. Port 22 of Unit 2 is connected by way of a cascade connection cable 35 to a port on a T-piece 34 which couples to the cascade module 33 of Unit 1 and this T-piece is connected by way of a cascade connection cable 36 to the cascade module 33 on Unit 0. The T-piece includes mulitplexers which are controlled in accordance with control signals sent between the T-piece and the cascade modules. Its essential function is to maintain the cascade

connection even if one or other of the units should be powered-down either deliberately or by virtue of failure.

Further information relating to the cascade connection, particularly for the switch type 4400 is available from GB patents GB-2365718 and GB-2369277. Other forms of cascade connection may be employed, such as the one described in GB patent application 0227048.6 filed 20 November 2002.

Figure 2 illustrates a typical router of the kind which is intended for use as a core router in the present invention. This is also a stackable device which can be stacked by means of architecture similar to that described in Figure 3.

The router unit 20 in Figure 2 has a multiplicity of ordinary or 'front panel' ports 21 and a 'cascade' port 22. The unit includes at least one and usually a multiplicity of (hardware) bridges or layer 2 switches 23. Each port 21 is connected to at least one of the bridges 23 and the or each cascade port 22 is connected to all the bridges or to a 'logical' internal port connected to all the bridges 23. The unit includes a router 24 which has at least two, and in the illustrated example three, router interfaces 25. Each router interface 25 is connected to one bridge only, although each bridge may be connected to more than one router interface 25. For each interface there is some means such as a register storing a MAC address and a network (IP) address for the interface. For controlling the bridges and the router there is a processor constituted by a CPU 26 which has recourse, by means of an appropriate memory system, to a management agent 27 and a routing protocol 28. The routing protocol controls routing tables 29. Also embedded in the unit, in for example an interface 30 for the management agent, are the unit's normal addresses, i.e. its MAC address its network IP address. These addresses are used for the management of the router, for example by an external network supervisor, and would according to prior practice by supplied by the CPU to the router interfaces.

Although it is not directly relevant to the present invention, a router of the kind shown in Figure 2 can be stacked and organised so that the stack has a lead router and subordinate routers in the manner described in GB patent application 0202425.5

Although it is, again, not fundamental to the present invention, routers and switches of the kind shown in Figure 1 can be made in a modular form wherein each of a multiplicity of groups of ports is provided on or associated with a respective module and the various modules are connected by high-speed links to all the other modules constituting the switch. The processing of a packet may be distributed between the modules so that for example look-ups are performed by the module associated with the port on which the packet is received but some final processing is made by the module having the egress port. One document which describes such organisation of a router is described in GB patent GB-2362289.

Description of packet switching according to the invention.

In a normal 'layer 3' router, all packets forwarded to the router are routed either to a specific destination if the IP address is known or to one of a multiplicity of default routers if the IP (network) destination address is unknown.

The invention has broadly two aspects. One is the provision of a new manner of organising the routing of packets at the edge of a network. A further aspect of the invention is the organisation of a switch for this purpose.

In particular, it is intended that a packet should be 'routed' locally in an edge switch if possible and the packet should be switched at the data link layer (layer 2) to a core router if it be not possible to route the packet locally. In effect the core router will be a default router but packets will be switched to it by means of layer 2 (media access control) switching rather than at the logical (layer 3) level.

Figure 4 illustrates part of a network organised according to the present invention and including an edge switch which is organised to act as a local router in accordance with the invention.

The network shown in Figure 5 includes a 'core' router 50, which may be a router organised on the lines of the router described with reference to Figure 2. The core router is coupled by an up-link 51 to an 'edge switch' constituted by a switch which is capable of layer 2 and layer 3 switching as described with reference to Figure 1. Ports on the edge router are coupled to a multiplicity of data terminal entities organised into a multiplicity of subnets; one of these subnets is shown as subnet 1 and includes a terminal PC1; another subnet is shown as subnet 2 and includes a terminal shown as PC2. Subnet 1 is regarded as being on 'VLAN 1' and devices on subnet 2 are defined as being on 'VLAN 2'. The core router may be coupled to other networks or subnets; it is shown as connected to a subnet 3 which includes a terminal shown as PC3.

Part of the database in the edge device, as shown in Figure 1, is a table of MAC addresses and corresponding destination ports. The table includes for each entry an additional bit field provided to indicate whether a received packet is to be subjected to a layer 3 look-up and switched accordingly, and this takes precedence over the destination port. The core router's MAC address is entered into this table so that all packets with this destination address will be forwarded to the layer 3 switch. This causes all packets destined for the core router to be sent to the layer 3 switch inside the edge router.

The various terminals will send ARP packets to determine where to send packets with a specified network address. If the ARP response comes from a remote network, it will have come via the router and will have the core router's MAC address as its MAC source address. The terminal can use this MAC address as the destination address for packets with this network (L3) address.

The layer 3 switching facility within the edge router will contain two types of routing information. This information may be in a single table or split into several tables.

The first type of routing information comprises the known routes. This is a list of all known destination addresses along with the information required to route the packet. The routing table is programmed with all the entries that are local to the edge switch. This information is obtained from the core router and in the example above this would be all the network addresses on subnet 1 and subnet 2. The source address entered in the routed packet should be the same as the source address of the core router.

A second type of routing information consists of default routes namely a list of routes which can be used if the packet does not match any of the entries in the routing table.

The embodiment to be described is implemented in hardware. The switch thus requires a mode to use the result of the layer 2 look-up if the layer 3 look-up fails. For this purpose no default routes would be programmed. If the layer 3 does not match any of the known routes, the packet is layer 2 switched using the result of the layer 2 look-up.

If on the other hand the invention is implemented in software, a default route table can be programmed with a default route that matches all packets. The source address to be inserted in the packet will be the MAC address of the edge switch and such packets will be routed to the core router.

Figures 5 and 6 illustrate the differences between ordinary layer 2 and layer 3 switching and the switching which is employed in the present invention.

In the ordinary scheme shown in Figure 5, a packet is received by the switch 2. A layer 2 address look-up is performed. It will be performed only if the destination address of the packet is in the same subnet as the source. The packet will be switched at layer 2 if an address match is found. The other possibility is that the layer 2 look-up result is to forward the packet to a layer 3 look-up. Here there are two possibilities. Either the lower 3 destination or the next hop is found, according to the routing tables, in which case packet L3 will be routed. If the look-up fails then the packet is passed to the CPU.

Figure 6 illustrates one implementation of the present invention, particularly suitable for a hardware version. The first stage is similar, in that the packet will be switched at layer 2 or will be forwarded to the layer 3 look-up. If the layer 3 destination is found by the layer 3 address look-up, then the packet will be routed. If however the look-up fails the packet will be switched using the previous result of a layer 2 look-up.

Figure 7 illustrates a basic network scenario. This resembles Figure 4 except that members of VLAN 1 such as PC4 are connected to the router 50. In such an implementation, a packet originating at PC1 and destined for PC2 on VLAN 2 will arrive at the edge switch 52 and will be routed locally and not forwarded to the router 50. A packet originating at PC1 on VLAN 1 and destined for PC3 on VLAN 3 will be routed to the router 50 and routed at that router to PC3. Packets originating at PC3 and destined for PC1 will be routed by the router 50 and then switched by the switch 52. Packets originating at or destined for PC1 and destined for or originating at PC4 as the case may be will be switched.

CLAIMS

1. A method of operating a network switch which is an edge switch in a packet-based communication network having a multiplicity of sub-nets, is arranged to receive and forward packets which include media access control address data and network address data, and is in communication with a core router, comprising:

performing a look-up in respect of a packet which is received by the edge switch from a source local to the edge switch and on a first sub-net and has a destination on a second sub-net;

forwarding the packet directly to its destination in response to the network address data in the packet, without the packet traversing the core router, when the destination is a local destination; and

forwarding the packet from the edge switch to the core router, in response to the media access control data in the packet, if the destination is not local to the edge switch;

said edge switch maintaining look-up tables of media access control addresses and network addresses for local sources and destinations on both the first and second sub-nets.

2. A network switch which is organized to receive and forward packets which include media access control address data and network address data and comprises:

means for performing a look-up in respect of a packet which is received by the edge switch from a source local to the edge switch and on a first sub-net and has a destination on a second sub-net;

means for forwarding the packet directly to its destination in response to the network address data in the packet when the destination is a local destination; and

means for forwarding the packet from the edge switch to core router's address, in response to the media access control data in the packet, if the destination is not local to the edge switch;

said switch having look-up tables of media access control addresses and network addresses for local sources and destinations on both the first and second sub-nets.

ABSTRACT

An edge switch maintains look-up tables of media access control addresses and network addresses for local sources and destinations at least two sub-nets, so that it can forward packets directly even when the local source and destination are on different sub-nets and forwards packets to a core router when the destination is not local to the edge switch.

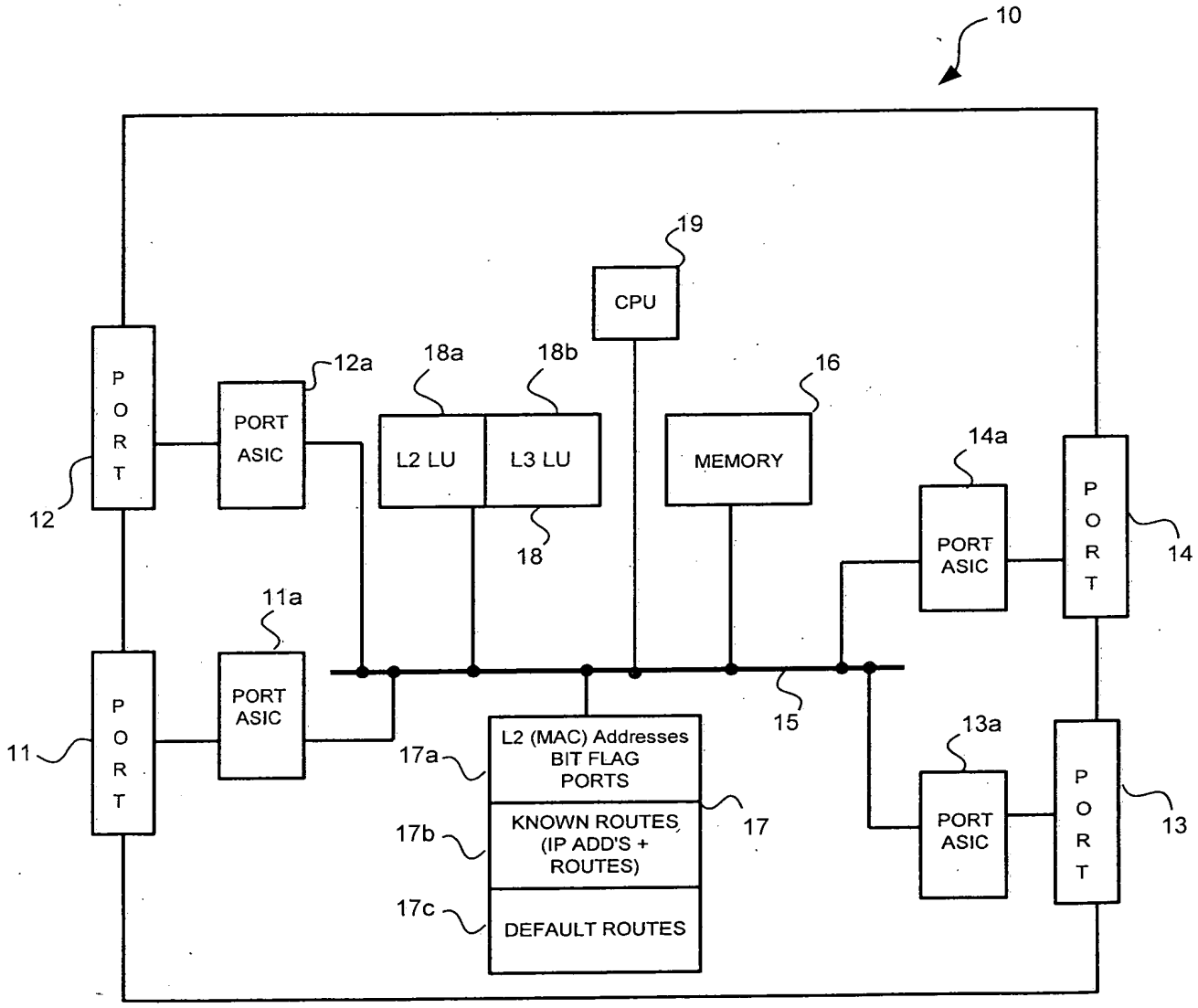
10
19
CPU
12a
18a
18b
16
P O R T
PORT ASIC
L2 LU
L3 LU
MEMORY
14a
12
18
PORT ASIC
P O R T
14
11a
P O R T
PORT ASIC
15
11
L2 (MAC) Addresses BIT FLAG PORTS
17a
13a
17b
KNOWN ROUTES (IP ADD'S + ROUTES)
17
PORT ASIC
P O R T
13
17c
DEFAULT ROUTES

FIG.1

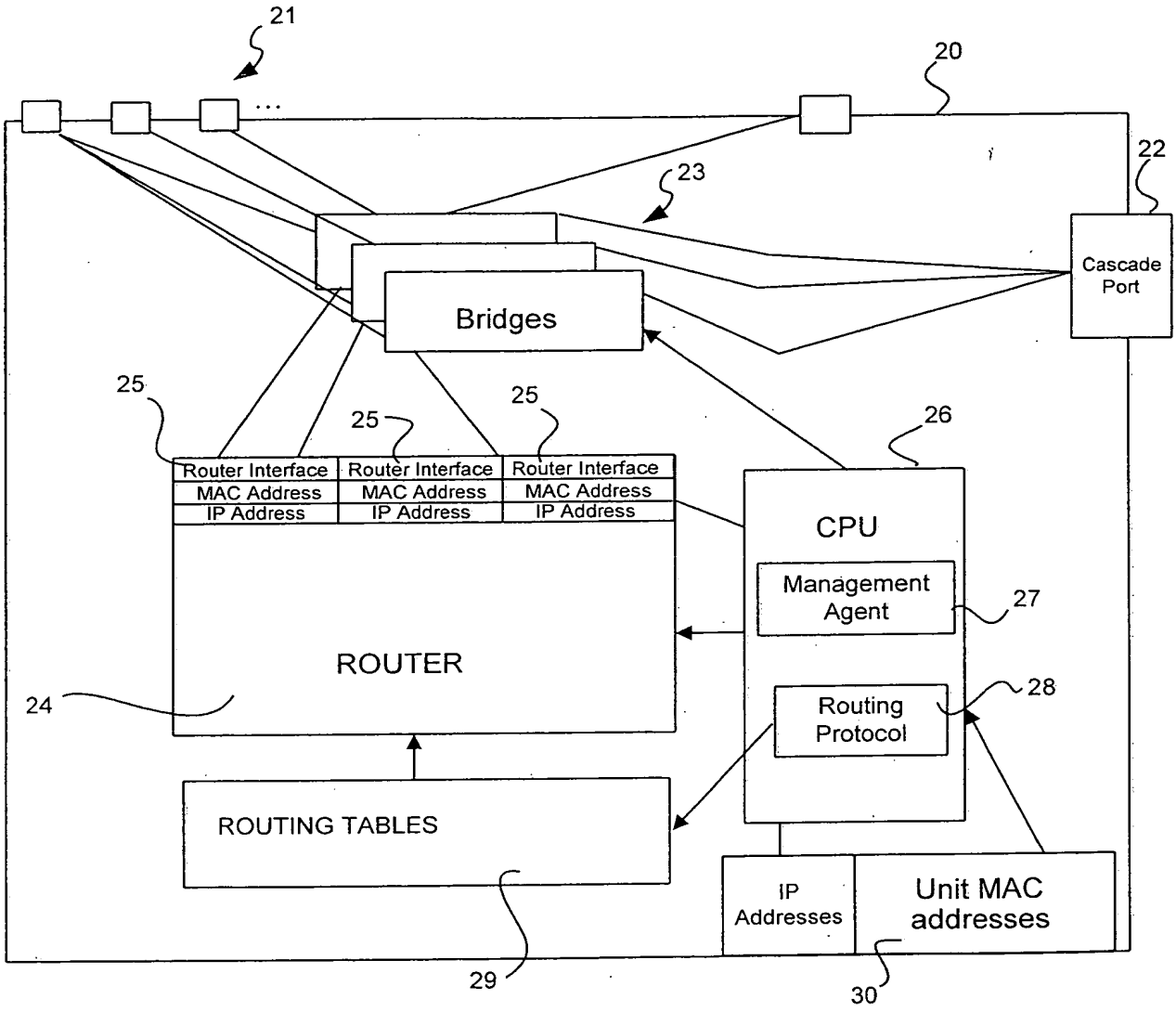
21
20
22
23
Cascade Port
Bridges
25
25
25
26
Router Interface
MAC Address
IP Address
Router Interface
MAC Address
IP Address
Router Interface
MAC Address
IP Address
CPU
Management Agent
27
ROUTER
24
28
Routing Protocol
ROUTING TABLES
IP Addresses
Unit MAC addresses
29
30

FIG.2

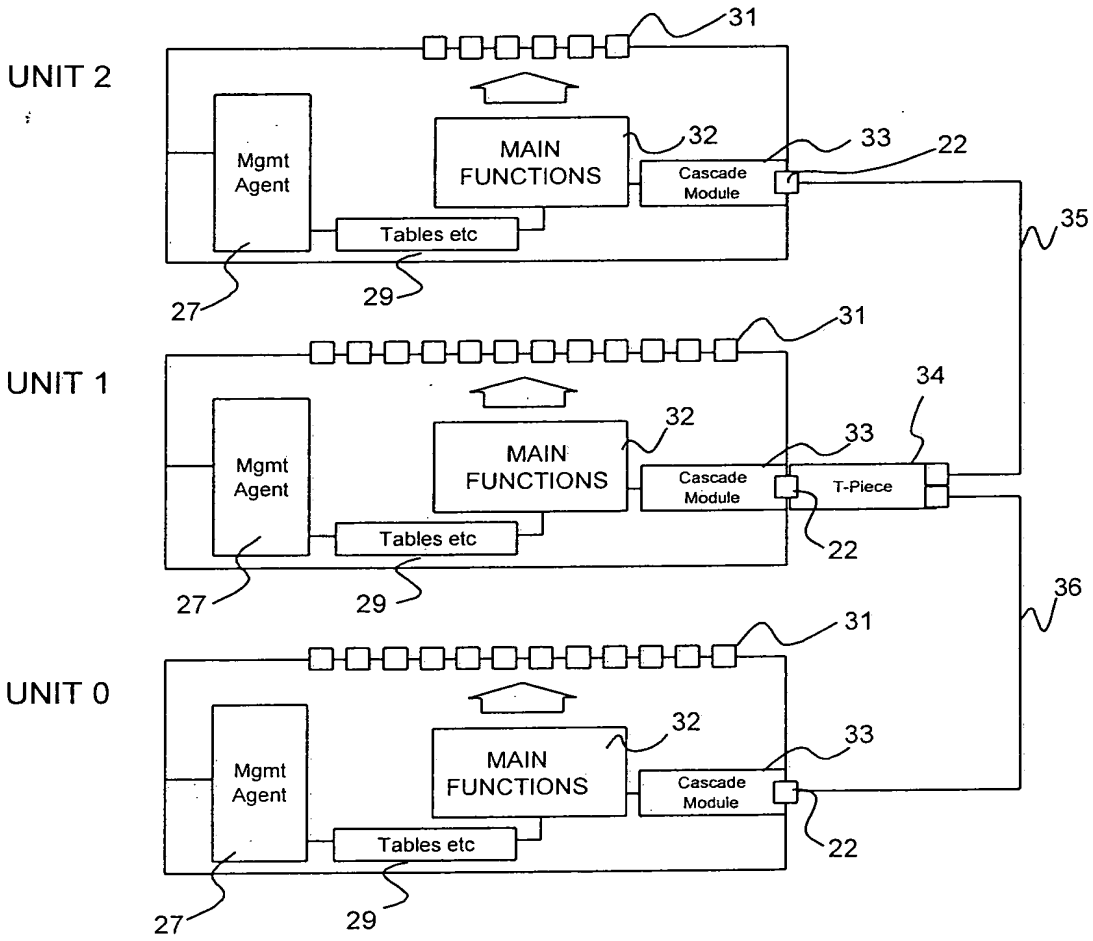
UNIT 2
UNIT 1
UNIT 0
Mgmt Agent
MAIN FUNCTIONS
Cascade Module
Tables etc
T-Piece
31
32
33
34
35
36
22
27
29

FIG.3

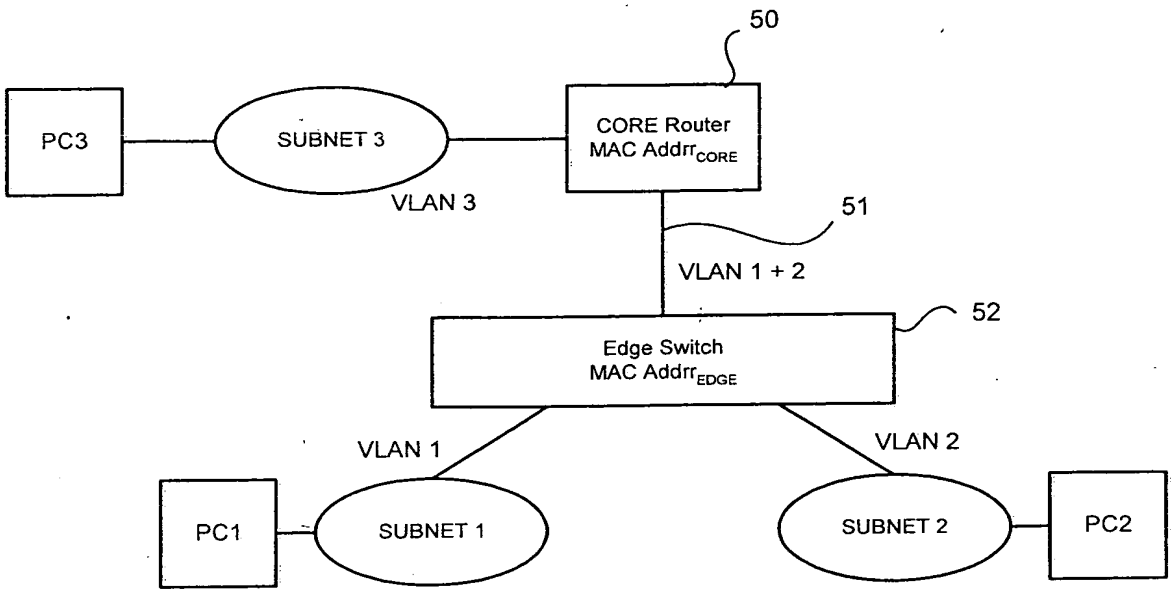
50
PC3
SUBNET 3
CORE Router
MAC Addrr$_{CORE}$
VLAN 3
51
VLAN 1 + 2
52
Edge Switch
MAC Addrr$_{EDGE}$
VLAN 1
VLAN 2
PC1
SUBNET 1
SUBNET 2
PC2

FIG.4

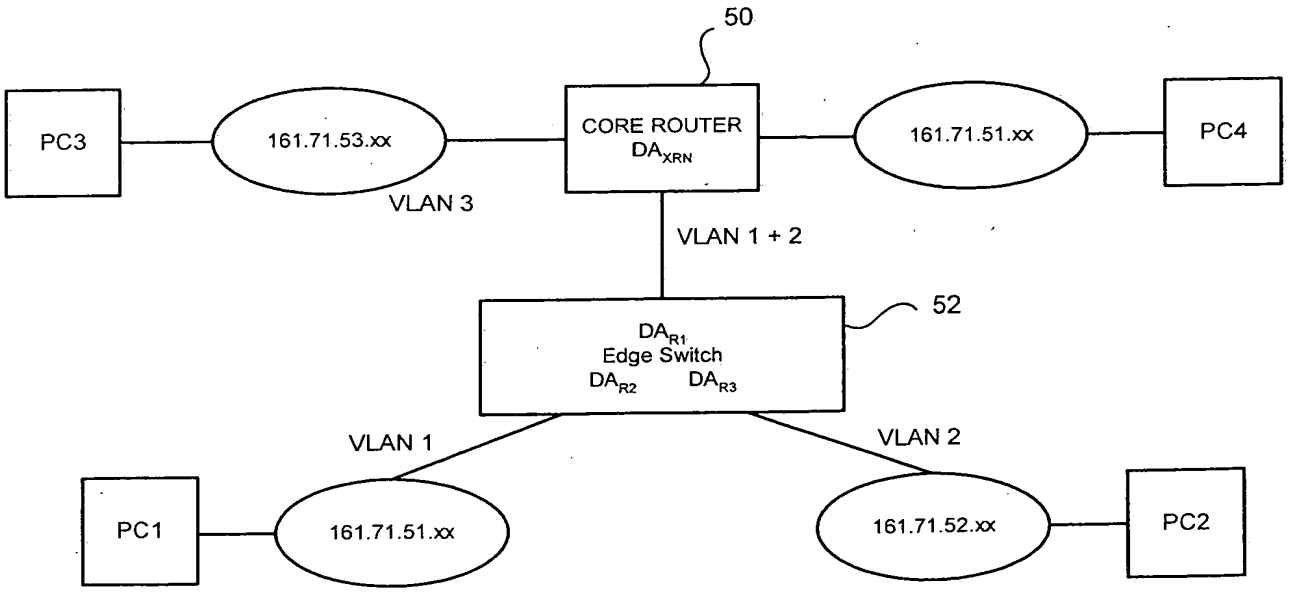
50
PC3
161.71.53.xx
CORE ROUTER
DA$_{XRN}$
161.71.51.xx
PC4
VLAN 3
VLAN 1 + 2
52
DA$_{R1}$
Edge Switch
DA$_{R2}$ DA$_{R3}$
VLAN 1
VLAN 2
PC1
161.71.51.xx
161.71.52.xx
PC2

FIG.7

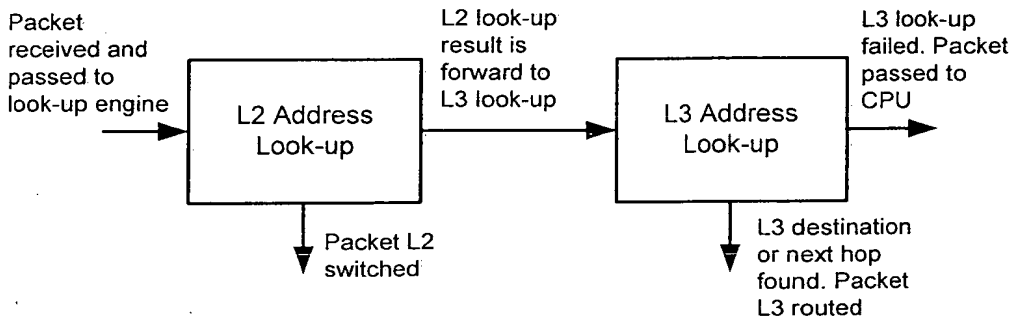
Packet received and passed to look-up engine
L2 Address Look-up
L2 look-up result is forward to L3 look-up
L3 Address Look-up
L3 look-up failed. Packet passed to CPU
Packet L2 switched
L3 destination or next hop found. Packet L3 routed

FIG.5

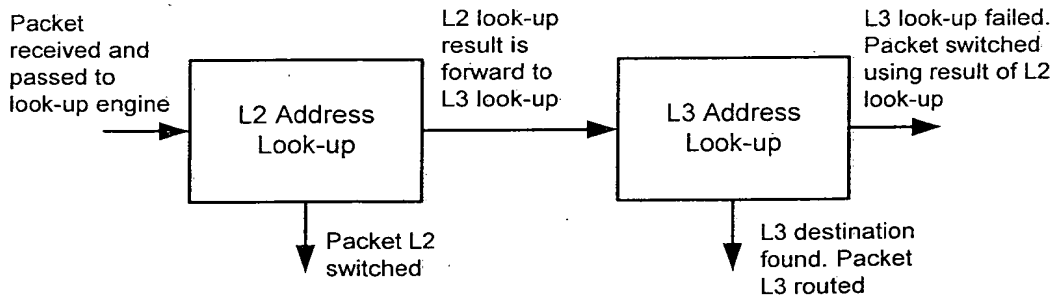
Packet received and passed to look-up engine
L2 Address Look-up
L2 look-up result is forward to L3 look-up
L3 Address Look-up
L3 look-up failed. Packet switched using result of L2 look-up
Packet L2 switched
L3 destination found. Packet L3 routed

FIG.6